United States Air Force
Air Force Materiel Command
Air Force Research Laboratory

The Isoazimuthal Perception of Sounds Across Distance: A Preliminary Investigation into the Location of the Audio Egocenter

Michael F. Neelon

University of Wisconsin Medical School
Madison WI 53705

Douglas S. Brungart
Brian D. Simpson

Air Force Research Laboratory

January 2005

Final Report for 1 October 2000 to 30 September 2004

20050421052

Human Effectiveness Directorate
Warfighter Interface Division
Battlespace Acoustics Branch
2610 Seventh Street
Wright-Patterson AFB OH 45433-7901

NOTICES

Please do not request copies of this report from the Air Force Research Laboratory. Additional copies may be purchased from:

> National Technical Information Service
> 5285 Port Royal Road
> Springfield, Virginia 22161

Federal Government agencies and their contractors registered with the Defense Technical Information Center should direct requests for copies of this report to:

> Defense Technical Information Center
> 8725 John J. Kingman Road, Suite 0944
> Ft. Belvoir, Virginia 22060-6218

TECHNICAL REVIEW AND APPROVAL

AFRL-HE-WP-JA-2005-0004

This report has been reviewed by the Office of Public Affairs (PA) and is releasable to the National Technical Information Service (NTIS). At NTIS, it will be available to the general public.

This technical report has been reviewed and is approved for publication.

FOR THE COMMANDER

//Signed//

STEPHEN D. STEELE
Deputy Chief, Warfighter Interface Division
Air Force Research Laboratory

REPORT DOCUMENTATION PAGE

1. REPORT DATE (DD-MM-YYYY)	2. REPORT TYPE	3. DATES COVERED (From - To)
January 2005	Final	1 Oct 2000 to 30 Sep 2004

4. TITLE AND SUBTITLE
The Isoazimuthal Perception of Sounds Across Distance: A Preliminary Investigation into the Location of the Audio Egocenter

5a. CONTRACT NUMBER

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER
61102F

6. AUTHOR(S)
Michael F. Neelon*, Douglas S. Brungart, Brian D. Simpson

*University of Wisconsin

5d. PROJECT NUMBER
2313

5e. TASK NUMBER
CB

5f. WORK UNIT NUMBER
72

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Air Force Research Laboratory, Human Effectiveness Directorate
Warfighter Interface Division
Battlespace Acoustics Branch
Air Force Materiel Command
Wright-Patterson AFB OH 45433-7901

8. PERFORMING ORGANIZATION REPORT NUMBER

AFRL-HE-WP-JA-2005-0004

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION / AVAILABILITY STATEMENT

Approved for public release; distribution is unlimited.

13. SUPPLEMENTARY NOTES
Originally published in The Journal of Neuroscience

14. ABSTRACT

Evidence indicates that both visual and auditory input may be represented in multiple frames of reference at different processing stages in the nervous system. Most models, however, have assumed that unimodal auditory input is first encoded in a head-centered reference form. The present work tested this conjecture by measuring the subjective auditory egocenter in six blindfolded listeners who were asked to match the perceived azimuths of sounds that were alternately played between a surrounding arc of far-field speakers and a hand-held point source located three different distances from the head. If unimodal auditory representation is head centered, then "isoazimuth" lines fitted to the matching estimates across distance should intersect near the midpoint of the interaural axis. For frontomedially arranged speakers, isoazimuth lines instead converged in front of the interaural axis for all listeners, often at a point between the two eyes. As far-field sources moved outside the visual field, however, the auditory egocenter location implied by the intersection of the isoazimuth lines retreated toward or even behind the interaural axis. Physiological and behavioral evidence is used to explain this change from an eye-centered to a head-centered auditory egocenter as a function of source laterality.

15. SUBJECT TERMS
auditory, egocenter, localization, multisensory, cortex visual

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE	SAR	9	Douglas S. Brungart
U	U	U			19b. TELEPHONE NUMBER (include area code) 937.255.3660

This page intentionally left blank.

Behavioral/Systems/Cognitive

The Isoazimuthal Perception of Sounds across Distance: A Preliminary Investigation into the Location of the Audio Egocenter

Michael F. Neelon,[1] Douglas S. Brungart,[2] and Brian D. Simpson[2]

[1]University of Wisconsin Medical School, Madison, Wisconsin 53705, and [2]Air Force Research Laboratory, Wright Patterson Air Force Base, Ohio 45433-7901

Evidence indicates that both visual and auditory input may be represented in multiple frames of reference at different processing stages in the nervous system. Most models, however, have assumed that unimodal auditory input is first encoded in a head-centered reference frame. The present work tested this conjecture by measuring the subjective auditory egocenter in six blindfolded listeners who were asked to match the perceived azimuths of sounds that were alternately played between a surrounding arc of far-field speakers and a hand-held point source located three different distances from the head. If unimodal auditory representation is head centered, then "isoazimuth" lines fitted to the matching estimates across distance should intersect near the midpoint of the interaural axis. For frontomedially arranged speakers, isoazimuth lines instead converged in front of the interaural axis for all listeners, often at a point between the two eyes. As far-field sources moved outside the visual field, however, the auditory egocenter location implied by the intersection of the isoazimuth lines retreated toward or even behind the interaural axis. Physiological and behavioral evidence is used to explain this change from an eye-centered to a head-centered auditory egocenter as a function of source laterality.

Key words: auditory; egocenter; localization; multisensory; cortex; visual

Introduction

Most spatial auditory research has assumed a head-centered coordinate system (Lewald and Ehrenstein, 1996; Stricanne et al., 1996; Duda and Martens, 1998; Jacobson et al., 2001) with its origin "halfway between the upper margins of the entrances to the two ear canals" (Blauert, 1983). However, little effort has been made to determine whether listeners judge the apparent locations of sounds relative to this interaural midpoint. In contrast, considerable research has been devoted to identifying the corresponding vantage point listeners use to judge the spatial locations of visual stimuli (Cox, 1999), often referred to as the visual "egocenter" (Roelofs, 1959).

Most methods for exploring the location of the visual egocenter have been based on Howard and Templeton's (1966) definition: "the location in the head toward which rods point when they are judged to be pointing directly to the self." Figure 1, A and B, illustrates two versions of this approach, in which egocenter estimates are obtained from the intersection of lines connecting visual objects at different distances in the same apparent direction. Current consensus is that the visual egocenter is located near or slightly behind the midpoint of the two eyes (Funaishi, 1926;

Komoda and Ono, 1974; Howard and Rogers, 1995; Cox, 1999). This suggests a discrepancy from the putatively head-centered auditory egocenter and implies a cross-modal mismatch between the apparent aural and visual locations of audiovisual stimuli close to the head.

Despite this possible discrepancy, we are aware of only one study that has attempted to empirically locate the auditory egocenter (Cox, 1999). That experiment used a variation of the approach used by Mitson et al. (1976) by replacing the distant visual targets with an arc of loudspeakers (see Fig. 1C). On each trial, a blindfolded listener adjusted the left–right position of a nearby vertical response handle to match the apparent direction of sound produced by one of the loudspeakers. Lines connecting the actual speaker locations to the apparent location judgments were extended back toward the head, and the auditory egocenter was then calculated from the centroid of their intersections. Results indicated that the egocenter was located near the back of the head (~12 cm behind the visual egocenter and 7 cm behind the interaural axis), suggesting the existence of large audiovisual parallax effects.

There are two possible methodological problems with this study, however. First, direction judgments were made with an unseen and unheard pointer that required listeners to transform perceived auditory locations into a kinesthetic frame of reference, potentially introducing error into the responses. More importantly, egocenter estimates were made from lines connecting actual loudspeaker locations with response locations, which assumed that target loudspeaker images were perceived at their true
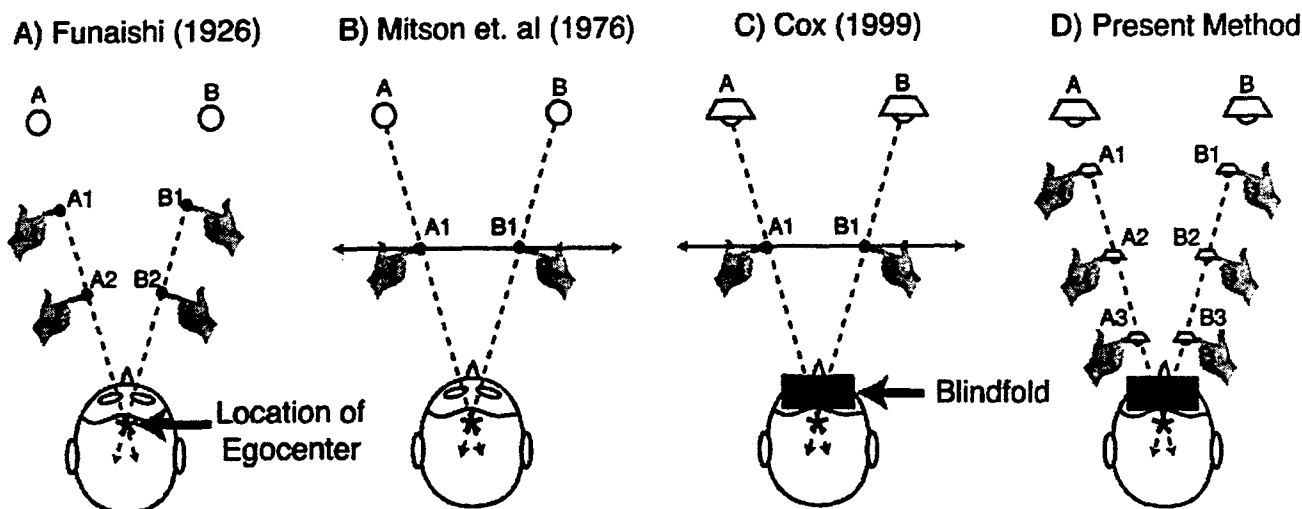
Figure 1. Methods used to estimate visual and auditory egocenters. *A*, From Funaishi (1926): observers matched the angle of a far visual target (*A, B*) at two distances (A1, A2 or B1, B2). Lines connecting far and near estimates for each target were extended back toward the head, and the visual egocenter was determined from their intersection (*). *B*, From Mitson et al. (1976) and Barbeito and Ono (1979): observers matched the angle of a far visual target using a track-mounted handle positioned at a single fixed distance in front of the head. Lines connecting the actual target locations (*A, B*) with the handle estimates (A1, B1) were extended back toward the head, and the egocenter was determined from their intersection. *C*, From Cox (1999): an auditory version of the method used by Mitson et al. (1976). For details, see Introduction. *D*, The present method is an auditory version of that used by Funaishi (1926) (*A*). For details, see Materials and Methods.

locations. If listeners mislocalized the loudspeakers, then the intersection lines would not represent lines of equal apparent azimuth, and the resulting auditory egocenter estimates would be invalid.

This paper describes a new attempt to measure the auditory egocenter using an adaptation of Funaishi's (1926) multiple response approach, in which listeners are required to make three matching responses for each fixed target location and the egocenter is estimated without reference to the actual target locations (see Fig. 1*A*). In the current study (see Fig. 1*D*), listeners move a nearby hand-held sound source to match the apparent locations of fixed target sounds, eliminating the need to translate the apparent audio locations of the target into a different modality.

Materials and Methods

Subjects. Six paid volunteer subjects (four male and two female) with clinically normal hearing and no previous experience with the procedures used in this experiment participated in the study.

Apparatus. The experiment was conducted in a medium-sized sound-treated hearing test chamber (4 × 4 × 4 m). The subjects were seated on a bench near the center of the chamber with their heads immobilized by a bite bar. Six small loudspeakers (Bose, Framingham, MA) were placed at eye level in an arc around the head (radius, 1.5 m), with speakers every 15° in azimuth from approximately −30° to the right to 45° to the left. Before each session, the subjects were blindfolded before being led into the test chamber and assisted onto the bench by the experimenter. This prevented them from seeing the physical arrangement of the speakers used in the experiment.

Once comfortably seated on the bench, subjects were handed a rigid "source wand" to manipulate the apparent location of a compact broadband sound source. The source itself consisted of an electromagnetic horn driver (DH1506; Electro-Voice, Burnsville, MN) connected to a long section of foam-covered flexible tygon tubing (internal diameter, 1.2 cm). This tube was acoustically terminated with a small piece of acoustic foam that was designed to minimize the occurrence of standing waves inside the source. The horn driver and most of the tubing were located on the floor in a corner of the test chamber and were acoustically isolated with sound-absorbent material. This acoustic tube source has the unique property that the sound it produces appears to originate from the opening at the end of the tube, which effectively acts as a compact, nondirectional, broadband acoustic point source (Brungart et al., 2000).

The last 60 cm of the tube was encased in a rigid polyvinyl chloride sleeve, which served as a "wand" that the subjects could easily use to control the location of the tip of the point source during the experiment.

The end of the source wand was equipped with an electromagnetic position sensor (FastTrak; Polhemus, Colchester, VT) that measured the location of the point source (i.e., the opening of the tube) during the experiment. The electromagnetic source for this position sensor was rigidly attached to the subject bench just under the subject's chin, and the location of the bench was clearly marked to ensure that its placement relative to the loudspeaker array was consistent across all of the trials of the experiment. This made it possible to accurately measure the absolute position of the point source relative to the six speakers in the fixed loudspeaker array.

Calibration. Before the start of each block of trials, a calibration procedure was used to determine the location and orientation of the subject's head relative to the fixed array of loudspeakers. In this procedure, the electromagnetic position sensor at the end of the source wand was used to measure three reference locations on the surface of the subject's bite bar-immobilized head: the opening of the left ear canal, the opening of the right ear canal, and the tip of the nose. These positions were used to define an egocentric spherical coordinate system, with its origin at the midpoint of the left and right ears, its "horizontal plane" defined by the locations of the left and right ears and the nose, and its median plane perpendicular to the interaural axis and passing as close as possible to the tip of the nose (Brungart et al., 2000). Within each session, all of the subject's responses were measured in this egocentrically defined coordinate system. The three positions were also used to measure the head width of each subject, as defined by the distance between the openings of the two ear canals.

These calibration measurements were used during subsequent data analyses to correct for any small changes in the relative locations of the fixed loudspeakers that might have occurred because of variations in subject placement on the bite bar across different experimental blocks. This correction was achieved by adjusting the responses within each block to compensate for the difference between the azimuthal orientation of the head within that block and the average azimuthal orientation of the head across all of the blocks collected for that subject. On the basis of these calibration measurements, the mean ± SD location of the speakers, in order from 1 to 6, averaged across all trials and all listeners were at the following angles relative to actual measured head orientations: −26.77 ± 1.28°, −11.97 ± 1.31°, 2.70 ± 1.36°, 16.84 ± 1.39°, 30.22 ±

1.39°, and 43.76 ± 1.37° (negative values indicate the listeners' right hemifield).

Procedure. Once the calibration procedure was complete, the experimenter left the test chamber and instructed the control computer to start data collection. Each trial of the experiment commenced with the onset of a continuous acoustic stimulus that alternated between one of the six loudspeakers in the fixed array and the acoustic point source at the end of the source wand held by the subject. Each stimulus presentation consisted of the following pattern: first, the fixed loudspeaker generated a 200 msec Gaussian noise burst; after a 100 msec interval of silence, the point source generated two 100 msec Gaussian noise bursts, separated by a 100 msec interval of silence; after another 100 msec interval of silence, the sequence started again with another 200 msec noise burst from the fixed loudspeaker. The Gaussian noise tokens were randomly selected, with replacement, from a set of 10 200 msec white-noise tokens and 10 100 msec white-noise tokens that were randomly generated at the start of each trial. These tokens were digitally low-pass filtered at 10 kHz, and the tokens' output to the point source was also filtered by a finite impulse response filter that was designed to match the frequency response of the point source at a 90° angle of incidence as closely as possible to the frequency response of one of the fixed loudspeakers at a direct angle of incidence.

After hearing this stimulus, the subject's task was to hold the hand-held point source vertically with its tip at the same level as the fixed loudspeaker and move it to a point where its apparent azimuth angle matched the apparent azimuth angle of the fixed sound source. Once the apparent directions of the two sources were matched, the subject responded by pressing a footswitch, which instructed the control computer to record the location of the point-source tip and randomly select another fixed loudspeaker location for the next trial. Spurious responses were reduced by preventing the listeners from responding until they heard the noise tokens alternate between the fixed loudspeaker and the point source at least four times. Although the subjects were allowed to manipulate the point source with either hand, most performed the matching task exclusively with their right (dominant) hand. [Handedness was assumed not to have influenced responses, because previous results in a similar task that required blindfolded listeners to move vertical handles to the perceived locations of sound sources showed no performance differences between the dominant and nondominant hand (Cox, 1999).]

Each experimental session consisted of three blocks of trials, with each block containing five repetitions at each of the six fixed loudspeaker locations. Before each block, the subjects were instructed to make their responses with the point-source wand at one of three distances: near, where they were instructed to hold the point source only a few inches from their head during the matching task; far, where they were told to hold the point source at arm's length during the matching task; and intermediate, where they were told to hold the source approximately halfway between these two extreme distances. Every session consisted of one block of trials in each of these three conditions, with the order randomized across subjects and across sessions.

In each of these three distance conditions, the far-field source was set to a comfortable listening level at the location of the listener (~70 dB sound pressure level), and the output level of the point source was scaled to maintain a similar level at the location of the listener. This required the point source to be attenuated by 0 dB in the far response blocks, 3 dB in the intermediate response blocks, and 6 dB in the near response blocks.

Each of the six subjects participated in a total of six sessions of the experiment. In each case, the data from the first session were discarded as training data, and only the data from the last five sessions were used for the data analysis. Thus, the data used in the analysis consisted of 25 matching estimates per point-source distance, per speaker, for a total of 450 matching estimates per subject.

Results

Initial assessment of point-source matching responses

The major goal of this work was to estimate the location of the auditory egocenter. However, such estimates are clearly influenced by how consistently the listeners wielded the point source

across distance in the azimuthal matching task. Reliability of the point-source matches was assessed by computing the grand average response SD at each point-source distance. Averaging across all listeners and speaker locations, the SDs for the three different point-source estimates were as follows: near, 4.22°; intermediate, 4.49°; and far, 6.02°. The range of listener-averaged SDs across the six speaker locations and three point-source distances was 3.78–7.45°. These values are in line with those reported by Makous and Middlebrooks (1990) for localization of frontal sources along the median horizontal plane, indicating that the current listeners were acceptably consistent in their matching judgments of apparent azimuth.

Isoazimuth lines and the auditory egocenter

Figure 2 applies the auditory Funaishi egocenter estimation method shown in Figure 1D to the location matching data collected in the experiment. Figure 2 presents a bird's-eye view of the individual calibration-corrected matching responses of each of the six subjects. In each panel of the figure, the listener's head (shown by a circle with a diameter equal to the average measured head size for that subject) is pointed toward positive values on the abscissa. The listener's left hemifield is denoted by positive ordinate values, and the right hemifield is denoted by negative ordinate values. The large numbered S's in the figure window show the locations of the six fixed loudspeakers relative to that listener's head. Each single response from an individual trial is represented by a number matching the far source for that trial. The ◊, □, and ○ symbols show the mean response locations for all of the near, intermediate, and far responses collected for a single fixed speaker location.

The six lines drawn in each panel of the figure represent linear fits of all the near, intermediate, and far matching responses collected for each of the six fixed loudspeaker locations. In other words, they represent the "isoazimuth" lines along which near, intermediate, and far sources all appeared to originate from the same direction relative to the listener. These isoazimuth lines were computed using a technique based on principal components analysis (Jackson, 1991), in which each line represents the first principal component extracted from all of the data points collected for a single fixed loudspeaker in the array. These first principal components accounted for almost all of the variability in the dataset (for all speakers and all subjects: mean, 97.47%; range, 95.56–99.73%).

The stars in each panel of the figure represent the estimated locations of the auditory egocenters, which were determined from the mean Cartesian coordinates of the 15 intersections that occurred between each pair of isoazimuth lines for each subject. The x and y locations of these mean egocenter estimates are also provided at the bottom left of each panel of the figure (along with the SE values in each dimension) and in Table 1. From these results, we note the following key points: (1) all six of the egocenter estimates fell very close to the median sagittal plane (range of mean y-axis values of egocenters, −0.65–1.10 cm), (2) the 95% confidence intervals of the x-axis values of the egocenter estimates for all six listeners fell in front of the geometric center of the head, (3) the average estimated egocenter location across all listeners (±1 SE) was $X = 6.1 ± 1.35$ cm in front of the interaural axis, $Y = 0.1 ± 0.27$ cm.

Thus, in contrast to the previous results by Cox (1999), this study found that the auditory egocenter is located very close to the generally accepted location of the visual egocenter (i.e., near the midpoint of the interocular axis). Furthermore, the results were remarkably consistent across the different listeners used in
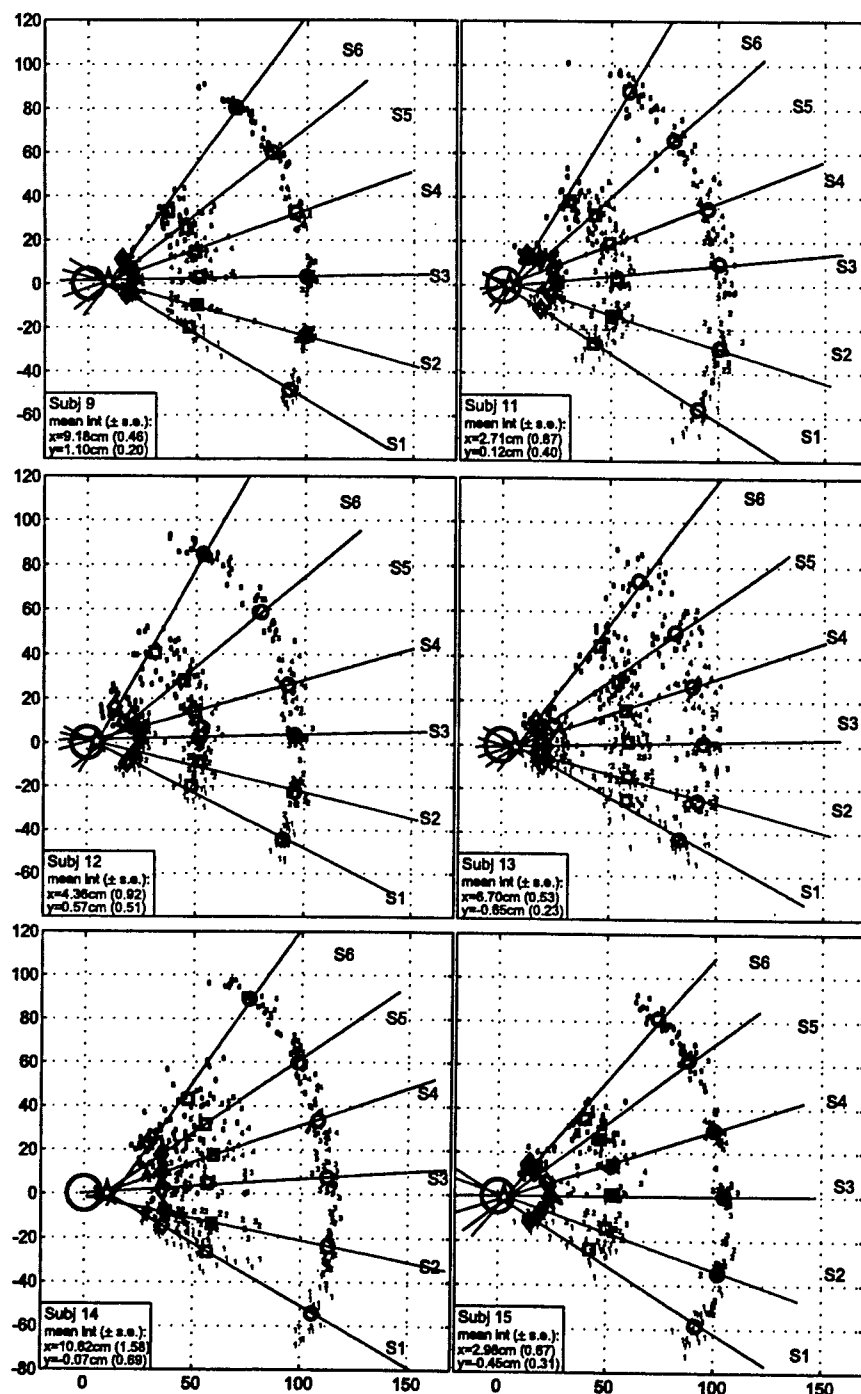
Figure 2. A bird's-eye view of the individual point-source estimates (small numerals) and fitted isoazimuth lines for the six listeners performing the azimuthal matching task to frontal speaker positions. The listener's head (large circle, averaged across blocks) is pointed toward positive values on the abscissa (all units are in centimeters). Symbols ◇, □, and ○ indicate mean point-source estimates for near, intermediate, and far distance placements, respectively. Also shown are mean speaker positions (S) and intersection (star) of isoazimuth lines. Insets display mean and SE values of the intersection in Cartesian coordinates.

tion is, in contrast to vision, an omnidirectional modality. Furthermore, there is less reason to expect an audiovisual parallax effect outside the visual field because it is impossible for listeners to simultaneously see and hear objects located in this region. Thus, there is no a priori reason to believe that the effective location of the auditory egocenter would need to be aligned with the visual egocenter for sources outside the field of view. In support of this conjecture, recent behavioral and physiological evidence suggests relatively less cross-modal interaction between audition and vision for multimodal sources located at extreme eccentricities, especially outside of the visual field (Linden et al., 1999; Falchier et al., 2002; Hairston et al., 2003). These arguments prompted a follow-up exploration of whether the auditory egocenter would remain unchanged as sound sources moved outside of the frontal binocular field of view (approximately ±60°) (Diffrient et al., 1981).

To explore this possibility, a replication of the experiment was conducted with a different arrangement of target speakers. This replication used the same basic setup and the same six listeners used previously, but it was conducted with the subject bench rotated counterclockwise relative to the speaker array. Thus, in the second experiment, the subject's medial sagittal plane bisected speaker locations 5 and 6, which were nominally located 30° and 45° to the left of the listener in the original experiment. Averaged across all trials and all listeners, the mean ± SD locations of the speakers (in order from 1 to 6) were at the following angles relative to actual measured head orientations: $-64.09 \pm 1.8°$, $-50.26 \pm 1.74°$, $-35.23 \pm 1.67°$, $-20.9 \pm 1.62°$, $-6.88 \pm 1.58°$, and $7.63 \pm 1.56°$ (in which negative values indicate locations in the listener's right hemifield).

For subject comfort, the number of trials per block was reduced from 30 to 18 (three repetitions at six speaker locations for a given point-source distance). Subjects completed at least six sessions in this replication, resulting in at least 18 matching estimates per point-source distance, per speaker. (Because of the rigors of sitting fixed to the bite bar, subject 9 was able to complete only three sessions for a total of nine estimates per distance per fixed loudspeaker location.) All other procedures and stimuli remained the same.

the experiment: all six listeners produced egocenter estimates in the same general vicinity, and most of them produced isoazimuth lines that intersected within a very tight spatial region near the front of the head.

One important limitation on the generality of the current data are that all of the near–far stimulus matching trials were conducted with target loudspeakers located in a 75° arc in front of the listeners. This arrangement fails to account for the fact that audi-

Auditory egocenter estimates for lateral source positions

Figure 3 presents a bird's-eye view of the individual matching responses and fitted isoazimuth lines for the new data, calculated using the same methods used in Figure 2. All orientations and symbols are the same as described in the previous figure. The

Table 1. Euclidean distances in centimeters of the mean isoazimuth line intersections (Fig. 2, star) from the interaural center point (0, 0) for each listener, with corresponding 95% confidence intervals

|	Subject 9	Subject 11	Subject 12	Subject 13	Subject 14	Subject 15
Mean x value of intersections	9.18	2.71	4.36	6.70	10.62	2.98
95% c.i. (x values)	[8.20, 10.17]	[0.84, 4.57]	[2.38, 6.33]	[5.57, 7.83]	[7.23, 14.01]	[1.54, 4.41]
(y values)	[0.68, 1.52]	[-0.73, 0.97]	[-0.52, 1.66]	[-1.14, -0.16]	[-1.55, 1.42]	[-1.11, 0.21]

c.i., Confidence interval.

fitted lines calculated from the first principal components again account for the variability in the matching responses quite well (for all speakers and all subjects: median, 98.29%; range, 96.26–99.78%).

There are several apparent differences between Figures 2 and 3. Most notably, mean isoazimuth intersections (stars) for four of the six listeners (all except 9 and 13) have shifted posteriorly relative to the intersections in the first experiment. These values, with corresponding 95% confidence intervals, are presented in Table 2. It is also clear from both Table 2 and Figure 3 that there is considerably more variation across the listeners in the fitted isoazimuth lines and the subsequent x- and y-coordinate values for the estimated egocenters. The six listeners generally appear to break into three groups regarding their egocenter confidence intervals: two listeners (9 and 13) produced egocenter estimates that were reliably in front of the geometric center of the head, two listeners (11 and 14) produced estimates that were not significantly different from the center of the head, and the two remaining listeners (12 and 15) produced estimates that reliably fell behind the center of the head.

Egocenter location as a function of source position

Superficially, the highly variable egocenter location estimates that occurred with the laterally placed speaker array results seem quite different from those measured for the frontally placed speaker array. However, a closer examination of the results suggests that the differences across the two experiments were primarily attributable to the extreme lateral speaker locations used in the second experiment rather than to changes in listeners' strategies or methodologies. First, estimates of response variability in the current experiment were similar to those found previously, suggesting that the listeners' perceptions of the far sources did not change in any qualitative way. Second, Figure 4 shows an analysis of the auditory egocenter similar to the ones used in Figures 2 and 3, which combined the data from all of the speaker locations in the two experiments that fell between −35 and 35° in azimuth (i.e., speakers 1–5 from the original experiment and speakers 3–6 from the replication). Although the average auditory egocenter
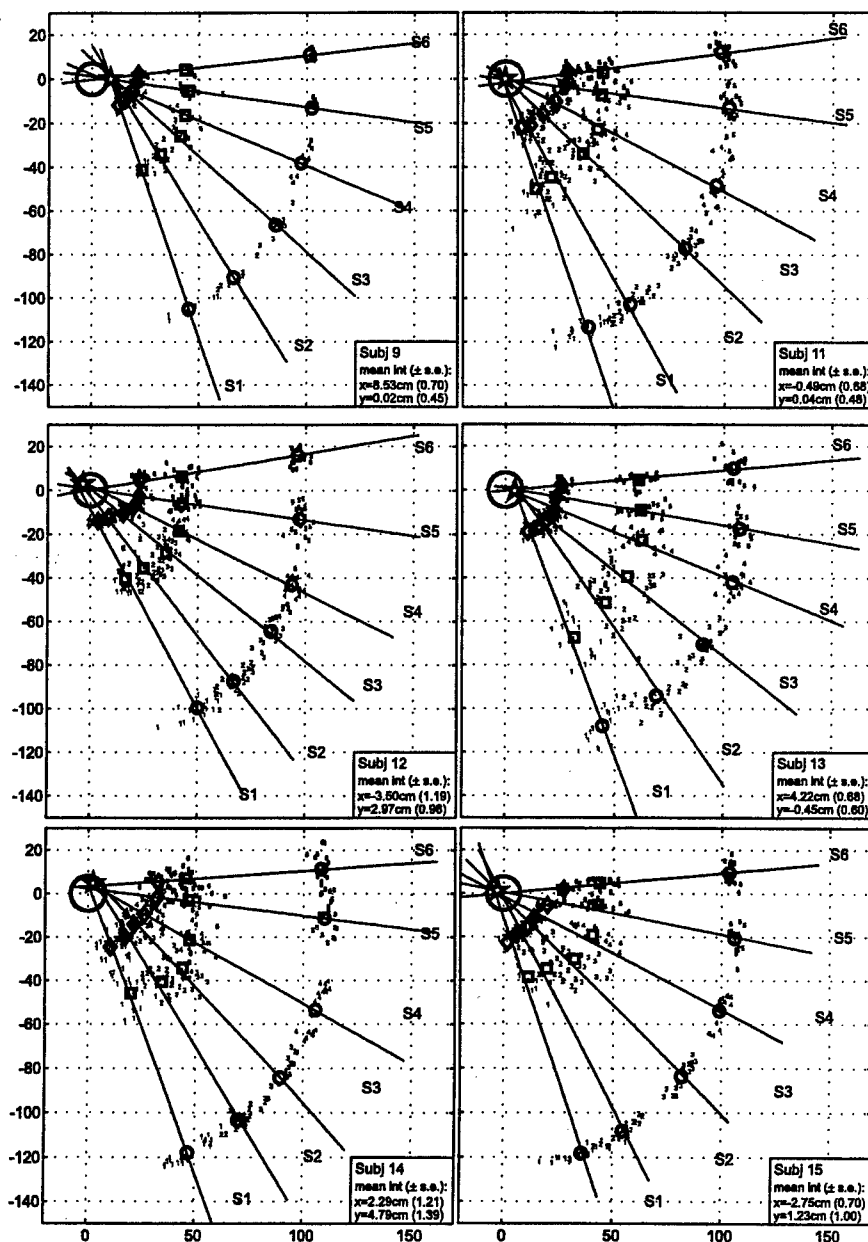
Figure 3. A bird's-eye view of the individual point-source estimates (points) and fitted isoazimuth lines for the six listeners performing the azimuthal matching task to lateral speaker positions. Orientation and symbols are the same as used in Figure 2.

estimated from the combined results was closer to the interaural axis than the average egocenter measured in the first experiment (x = 4.3 vs 6.1 cm), all of the subjects again had mean egocenter locations that fell in the front half of the head.

These results suggest that the differences in the egocenter locations found between the two experiments were the product of differences in speaker locations rather than any underlying variability in the egocenter estimation methods used. More specifically,

5

Table 2. Euclidean distances in centimeters of the mean isoazimuth line intersections (Fig. 3, star) from the interaural center point (0, 0) for each listener in the follow-up experiment, with corresponding 95% confidence intervals

	Subject 9	Subject 11	Subject 12	Subject 13	Subject 14	Subject 15
Mean x value of intersections	8.53 cm	−0.49	−3.50	4.22	2.29	−2.75
95% c.i. (x values)	[7.03, 10.03]	[−1.96, 0.98]	[−6.06, −0.94]	[2.77, 5.67]	[−0.32, 4.90]	[−4.25, −1.25]
(y values)	[−0.95, 0.99]	[−1.00, 1.08]	[0.91, 5.03]	[−1.73, 0.83]	[1.81, 7.77]	[−0.91, 3.37]

c.i., Confidence interval.
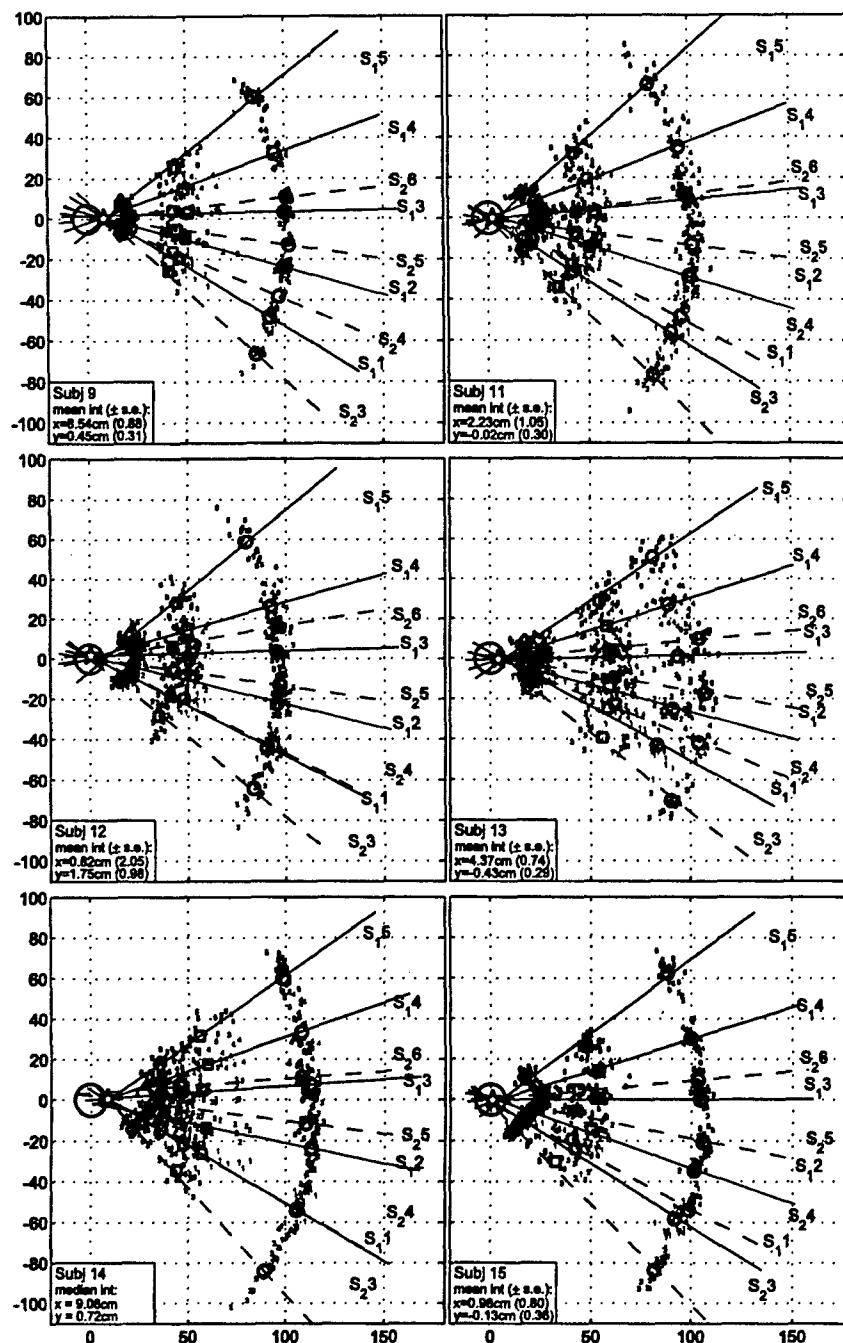
Figure 4. A composite bird's-eye view of the individual point-source estimates (small numerals) and fitted isoazimuth lines for all azimuthal matching data to speakers encompassing the frontal visual field from both the original and replication experiments (speaker numbers 1–5 and 3–6, respectively). Head position (large circle) is averaged from measurements for each listener across the two experiments. Orientation and all other symbols are the same as used in Figures 2 and 3. For subject 14, the median isoazimuth intersection was used to estimate the auditory egocenter because isoazimuth lines for midline sources resulted in a highly skewed mean egocenter estimate.

it indicates that the effective location of the auditory egocenter for lateral sound sources is more variable across different subjects and is located further toward the back of the head on average. This shift in the composite auditory egocenter from in front of to behind the interaural axis is discussed in more detail in the Discussion.

Discussion

Although most auditory models have assumed that spatial auditory information is encoded in head-centered coordinates, relatively little effort has been made to validate this conjecture experimentally. [For auditory midline estimates under headphones, see Lewald and Ehrenstein (1996).] In this experiment, auditory egocenter locations were estimated from isoazimuth lines that connected near, intermediate, and far listener estimates of the same apparent auditory angles. For frontal sources (±30° around midline), the results suggest an auditory egocenter located slightly in front of the interaural axis in the median sagittal plane, a point approximately corresponding to the accepted location of the visual egocenter (near the midpoint of the line connecting the two eyes). However, as sound sources move outside the frontal binocular visual field (beyond approximately ±60°), the auditory egocenter shifts posteriorly for some listeners. This direction-dependent shift in egocenter appears to be a reliable shift in the isoazimuthal perception of sounds across distance for these listeners.

Psychophysical and physiological foundations for frontal auditory egocenters

A number of studies provide indirect evidence to support these findings of an anterior auditory egocenter for frontomedial sources and a posterior egocenter for sources outside the visual field. The anterior auditory egocenter location that occurs for frontal sources might be directly related to the interaural time difference (ITD) and interaural level difference (ILD) cues that dominate the perceived horizontal locations of sounds (Grantham, 1995). As nearby sounds approach the head, there is generally a large increase in the ILD but only a modest increase in the ITD (Duda
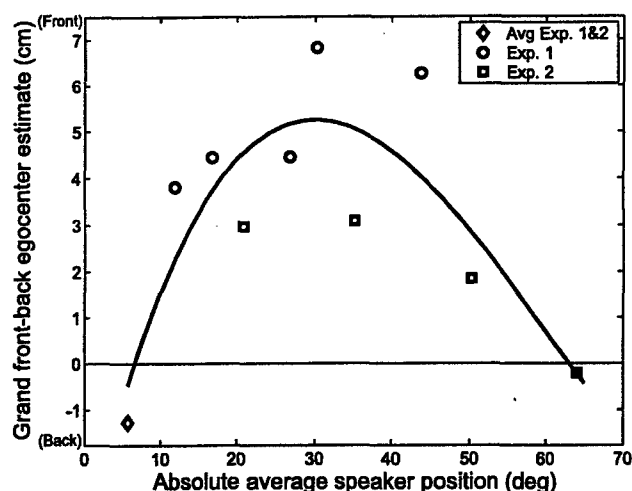
Figure 5. The intersection points of each composite isoazimuth line with the median sagittal line (x, 0) as a function of absolute average speaker position for the data from the two experiments (symbols) combined across all listeners. "Front" or "Back" indicates in front of or behind the composite interaural axis (solid line). The diamond indicates the mean of the intersection points for all speakers less than ±10°.

and Martens, 1998; Brungart and Rabinowitz, 1999; Shin-Cunningham et al., 2000). This may cause listeners who weight ILD more heavily than ITD in judgments of apparent azimuth (Yost, 1981; Dye et al., 1994; Hartmann, 1997; Altman et al., 1999) to perceive near-field medial sources and lateral far-field sources at the same apparent azimuth locations, thus causing an anterior shift in the effective location of the auditory egocenter similar to that seen for frontal sources in these experiments. By the same token, listeners using different interaural weighting schemes (Dye et al., 1994; Hartmann, 1997) may be more or less prone to exhibit anteriorly shifted auditory egocenters, which could explain some of the variability in the estimates reported here.

An extensive body of single-cell recording studies may also provide neurophysiological evidence that the audio and visual frames of reference can be aligned for stimuli inside the observer's field of view. Neurons in the superior colliculus and its associated cortical regions appear to be involved in transforming auditory information from an initial craniocentric representation into the retinocentric frame of reference needed to make orientation responses (Sparks and Nelson, 1987; Russo and Bruce, 1994). Stricanne et al. (1996) have further found acoustically responsive cells in the lateral intraparietal (LIP) area that characterize space in eye-centered, head-centered, and intermediate coordinate systems. This result suggests that listeners might represent auditory information in several egocentric coordinate systems, which could explain why some listeners in this experiment consistently exhibited anterior auditory egocenters, whereas others exhibited posteriorly shifted egocenters as sound sources moved outside the visual field.

Nonvisual cortical influence on peripheral visual cortex may predict changes in auditory egocenter location

The angle-dependent changes in the effective auditory egocenter locations that were exhibited by listeners in this study did not change monotonically. Rather, egocenters reached their most anterior positions for sources around ±30° and then retreated to more posterior positions for sources outside this range. This trend is visualized in Figure 5, which plots the intersection of each

composite isoazimuth line with the median sagittal line (x, 0) as a function of speaker position. These composite lines were estimated from the first principal components extracted from the entire set of point-source responses for each speaker combined across all listeners in both experiments. Because the shallow isoazimuth lines that occur for sources near the midline result in more variable intersection estimates, the mean is taken for isoazimuth line intersections with the sagittal line for speakers less than ±10° (diamond) and represented as a single point in Figure 5.

A third-order polynomial fitted to the data (solid line) clearly shows the nonmonotonic trend of the auditory egocenter estimates along the sagittal axis. Several points are worth noting about this result. First, the egocenter is estimated to be near or slightly behind the center of the head for averaged sources around midline. This finding can be predicted from the fact that no auditory parallax effect should theoretically arise for sources directly at (0,0) (indeed, the sagittal location of the auditory egocenter cannot be estimated for such sources). In fact, the slightly posterior data value for averaged medial sources plotted by the diamond in Figure 5 is approximately consistent with the posterior egocenter estimates reported by Cox (1999) for data that were primarily collected using sources at ±15°. Estimates formed from these source positions may thus have skewed Cox's results to more posterior values (e.g., extreme posterior egocenter estimates for medial sources can be seen in the present results in Fig. 4, subject 14).

Second, the fitted egocenter function reaches its peak frontal values for sources near 30° and then declines to near zero for sources near 60°. This frontal peak of the auditory egocenter for near medial sound sources, and later retreat for peripheral sources outside the binocular visual range, may have a physiological basis. A recent anatomical study in the monkey has found neural projections from auditory cortex to areas in visual cortex subserving peripheral visual fields (Falchier et al., 2002). These projections appear minimal for visual cells responding to medial sources (near 0°) but increase exponentially for visual cells responding to eccentricities of 15–20°. One explanation of such connections is that auditory influence on visual perception should be strongest for near eccentric stimuli to assist orienting behavior. (Stimuli at midline are more likely to be already foveated and thus may not require additional orientating responses.) If reciprocal connections exist, then visual influence on the auditory egocenter may likewise be strongest for slightly eccentric stimuli.

Because Falchier et al. (2002) did not measure corticocortical connections for visual neurons responding to stimuli more peripherally than 20°, it is uncertain whether the strong auditory cortical connections increase or decrease for more extreme visual eccentricities. An explanation offered here is that, as sound sources move out of binocular view, successful orientation must engage proportionally more head and body movements. This suggests that a craniocentric auditory egocenter may be usefully invoked for extreme eccentricities outside the visual field. Such changes in the relative amount of eye versus head movements to auditory targets as a function of source laterality have been reported in human listeners by Goldring et al. (1996). This hypothesis of audiovisual change in the pursuit of orientation would predict well the nonmonotonic frontal egocenter trend seen in Figure 5 and is further supported by psychophysical studies that have shown that audiovisual influences appear to decrease as sound sources move toward extreme eccentricities (Lewald and Ehrenstein, 1998; Hairston et al., 2003). Altogether, these data support a model in which the auditory egocenter is eye centered for

frontomedial stimuli but moves to a more craniocentric frame of reference as sound sources outside the binocular visual field.

Final methodological considerations

Although every possible effort was made to control extraneous influences in these experiments, some nonperceptual factors may still have contributed to the results reported here. First, no feedback was given to any subjects in the matching task beyond the initial instructions. Given that head-, eye-, and intermediate-centered cells could simultaneously be found in a single listener, it is possible that the lack of feedback allowed individual listeners to make matching judgments on each trial as they saw fit. In an attempt at consistency, listeners may have tacitly focused on and amplified a single perceptual coordinate system.

Another source of variability may have been the influence of eye position on the perceived locations of the sources. Although listeners' head positions were constrained through use of a bite bar, eye positions were not controlled in these experiments. Several psychophysical experiments have reported an effect of varying eye position on auditory lateralization and localization (Weerts and Thurlow, 1971; Rakerd and Hartmann, 1985; Lewald and Ehrenstein, 1996, 1998), which may have impacted the auditory egocenter measurements.

Ideally, an experiment would be conducted to extend the analysis of the egocenter to sources in all directions. Rear sources could help distinguish whether differences in location of the egocenter across individual listeners are the result of a reliance on interaural differences rather than actual localization. The current experimental design is most likely not feasible for collecting such data, however, because matching sources across distance using a hand-held source wand is impractical for rear sources.

Finally, the use of a hand-held wand may have involved multimodal cortical areas such as LIP in the azimuthal matching task. The transformation of auditory information into an eye-centered frame of reference in such areas may have potentially amplified the frontal auditory egocenters reported here. Many of these physical limitations could be overcome by conducting a similar experiment using virtual rather than free-field sound sources. Preliminary data have been collected in our laboratory in which listeners adjusted the position of a nearby virtual source to match the apparent azimuth of a more distant virtual source (Brungart et al., 2002). Localization and distance cues for the stimuli were created using nonindividualized head-related transfer functions (HRTFs). Data for three subjects performing this task produced oculocentric egocenters consistent with those found in the first experiment reported here. Ultimately, this virtual experiment should be reproduced using individualized HRTFs to avoid any influence of nonveridical spatial acoustic information on the matching task.

References

Altman JA, Variaguina OV, Nikitin NI (1999) Lateralization of a moving auditory image: interrelation of interaural time and intensity differences. J Acoust Soc Am 105:366–376.

Barbeito R, Ono H (1979) Four methods of locating the egocenter: a comparison of their predictive validities and reliabilities. Behav Res Methods Instr 11:31–36.

Blauert J (1983) Spatial hearing—the psychacoustics of human sound localization, p 14. Cambridge, MA: MIT.

Brungart D, Rabinowitz W (1999) Auditory localization of nearby sources. I. Head related transfer functions. J Acoust Soc Am 107:2224–2235.

Brungart DS, Durlach NI, Rabinowitz WM (2000) Evaluation of response methods for the localization of nearby objects. Percept Psychophys 62:48–65.

Brungart DS, Neelon M, Kordik A, Simpson B (2002) A triangulation method for determining the perceptual center of the head for auditory stimuli. Proceedings of the EAA Forum Acusticum, Sevilla, Spain, September.

Cox PH (1999) An initial investigation of the auditory egocenter: evidence for a "cyclopean ear." PhD thesis, North Carolina State University.

Diffrient N, Tilley AR, Harmon D (1981) Humanscale 4/5/6, Sec 6a. Cambridge, MA: MIT.

Duda RO, Martens WL (1998) Range dependence of the response of a spherical head model. J Acoust Soc Am 104:3048–3057.

Dye RH, Yost WA, Stellmack MA, Sheft S (1994) Stimulus classification procedure for assessing the extent to which binaural processing is spectral analytic or synthetic. J Acoust Soc Am 96:2720–2730.

Falchier A, Clavagnier S, Barone P, Kennedy H (2002) Anatomical evidence of multimodal integration in primate striate cortex. J Neurosci 22:5749–5759.

Funaishi S (1926) Uber das zentrum der sehrichtungen. Albrect von Graefes Archiv fur Ophthalmalogie 117:296–303.

Goldring JE, Dorris MC, Cornell BD, Ballantyne PA, Munoz DP (1996) Combined eyehead gaze shifts to visual and auditory targets in humans. Exp Brain Res 111:68–78.

Grantham DW (1995) Spatial hearing and related phenomena. In: Hearing (Moore BCJ, ed), pp 297–339. New York: Academic.

Hairston WD, Wallace MT, Vaughan JW, Stein BE, Norris JL, Schirillo JA (2003) Visual localization ability influences cross-modal bias. J Cognit Neurosci 15:20–29.

Hartmann WH (1997) Listening in a room and the precedence effect. In: Binaural and spatial hearing in real and virtual environments (Gilkey RH, Anderson TR, eds), pp 191–210. Mahwah, NJ: Erlbaum.

Howard I, Templeton W (1966) Human spatial orientation, Chap 11, p 274. London: Wiley.

Howard IP, Rogers BJ (1995) Binocular vision and stereopsis, Chap 14. New York: Oxford UP.

Jackson JE (1991) A user's guide to principal components, pp 1–25. New York: Wiley.

Jacobson G, Poganiatz I, Nelken I (2001) Synthesizing spatially complex sound in virtual space: an accurate offline algorithm. J Neurosci Methods 106:29–38.

Komoda MK, Ono H (1974) Oculomotor adjustments and size-distance perception. Percept Psychophys 15:353–360.

Lewald J, Ehrenstein WH (1996) The effect of eye position on auditory lateralization. Exp Brain Res 108:473–485.

Lewald J, Ehrenstein WH (1998) Auditory-visual spatial integration: a new psychophysical approach using laser pointing to acoustic targets. J Acoust Soc Am 104:1586–1597.

Linden JF, Grunewald A, Andersen RA (1999) Responses to auditory stimuli in macaque lateral intraparietal area. II. Behavioral modulation. I Neurophysiol 82:343–358.

Makous JC, Middlebrooks JC (1990) Two-dimensional sound localization by human listeners. J Acoust Soc Am 87:2188–2200.

Mitson GL, Ono H, Barbeito R (1976) Three methods of measuring the location of the egocenter: their reliability, comparative locations and intercorrelations. Can J Psychol 30:1–8.

Rakerd B, Hartmann WM (1985) Localization of sound in rooms. II. The effect of a single reflecting surface. J Acoust Soc Am 78:524–533.

Roelofs CO (1959) Considerations on the visual egocenter. Acta Psychologica 16:226–234.

Russo GS, Bruce CJ (1994) Frontal eye field activity preceding aurally guided saccades. J Neurophysiol 71:1250–1253.

Shinn-Cunningham B, Santarelli S, Kopco N (2000) Tori of confusion: binaural localization cues for sources within reach of a listener. J Acoust Soc Am 107:1627–1636.

Sparks DL, Nelson JS (1987) Sensory and motor maps in the mammalian superior colliculus. Trends Neurosci 10:312–317.

Stricanne B, Andersen RA, Mazzoni P (1996) Eye-centered, head-centered and intermediate coding of remembered sound locations in area LIP. J Neurophysiol 76:2071–2076.

Weerts TC, Thurlow WR (1971) The effects of eye position and expectation on sound localization. Percept Psychophys 9:35–39.

Yost WA (1981) Lateral position of sinusoids presented with interaural intensive and temporal differences. J Acoust Soc Am 70:397–409.